Theor Appl Genet (1996) 92:957–965

S. Grandillo · S. D. Tanksley

Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium*

Received: 22 May 1995 / Accepted: 22 September 1995

Abstract A population of 257 BC_1 plants was developed from a cross between an elite processing line of tomato (*Lycopersicon esculentum* cv'M82-1-7') and the closely related wild species *L. pimpinellifolium* (LA1589). The population was used to construct a genetic linkage map suitable for quantitative trait locus (QTL) analysis to be conducted in different backcross generations. The map comprises 115 RFLP, 3 RAPD and 2 morphological markers that span 1279 cM of the tomato genome with an average distance between markers of 10.7 cM. This map is comparable in length to that of the high-density RFLP map derived from a *L. esculentum* × *L. pennellii* F_2 population. The order of the markers in the two maps is also in good agreement, however there are considerable differences in the distribution of recombination along the chromosomes. The segregation of six GATA-containing loci and 47 RAPD markers was also analyzed in subsets of the population. All of the microsatellite loci and 35 (75%) of the RAPDs mapped to clusters associated with centromeric regions.

Key words Mapping · RFLPs · GATA microsatellites · RAPDs · *L. pimpinellifolium*

Introduction

Molecular mapping of crop plant genomes has advanced dramatically in the past 10 years. Currently, molecular linkage maps are available for most major crops and many minor crops (O'Brien 1993; Chittenden et al. 1994; Reinish et al. 1994), and work is starting for several other plant species (Faure et al. 1993, banana; Kennard et al. 1994, cucumber; Yang and Quiros 1995, celery).

Communicated by M. Koorneef

S. Grandillo · S. D. Tanksley (✉)
Department of Plant Breeding and Biometry, 252 Emerson Hall, Cornell University, Ithaca, NY 14853-1902, USA

These maps and their associated technology have been used successfully for a number of applications in plant breeding and genetics including: (1) characterization of genetic variation in germplasm collections (Figdore et al. 1988; Gawel et al. 1992), (2) gene tagging (i.e. identification of markers tightly linked to major genes) (Schüeller et al. 1992; MacKill et al. 1993), (3) map-based gene cloning (Arondel et al. 1992; Martin et al. 1993), and (4) analysis of quantitative traits (Edwards et al. 1987; Paterson et al. 1988, 1991; Tanksley 1993). The analysis of quantitative traits is especially interesting for plant breeding in that it opens the door to a more precise manipulation of quantitative traits such as yield, nutritional quality and flavor. However, for quantitative trait locus (QTL) analysis to be performed, a molecular map which provides full covergae of the genome of interest is sought.

The molecular markers most commonly used are restriction fragment length polymorphisms (RFLPs) (Botstein et al. 1980). Despite the many advantages offered by this class of markers, their use remains laborious and the level of polymorphism can become limiting, especially for crops with a narrow genetic base such as cotton, soybean and tomato. Random amplified polymorphic DNAs (RAPDs) are simpler to assay than RFLPs and can detect polymorphisms in both low-copy and repetitive DNA sequences (Williams et al. 1990). One limitation of this class of markers is that most RAPD markers are dominant and can detect only two alleles of a locus (presence or absence of the marker). Therefore, dominant RAPD markers provide less genetic information than RFLPs in certain segregating populations. RFLPs and RAPDs have been used in combination to construct a genetic map of *Arabidopsis thaliana* (Reiter et al. 1992). Integrated linkage maps based on both types of markers are now being generated for other crops (Rajapakse et al. 1995, peach; Cai et al. 1994, *citrus*).

An alternative source of DNA polymorphism has recently been described based on variation in the length of simple sequence repeats (SSRs) (also called micro-

satellites) (Tautz 1989; Weber and May 1989). SSRs are relatively short stretched of DNA that consist of tandemly repeated sequence motifs 1–4 bp in length. This class of markers has the advantages of showing high polymorphism and being suitable for polymerase chain reaction (PCR) amplification. In humans, $(CA)_n$ repeats are widely distributed at a high frequency throughout the genome and have therefore been used for constructing a saturated genetic map (Weissenbach et al. 1992). In other mammals, such as the mouse, SSRs have also been successfully used to construct saturated linkage maps (Dietrich et al. 1992). Tetranucleotide simple sequence repeats from the classes GACA have, in contrast, been shown to be clustered mainly on the short arms of all acrocentric autosomes in the human genome and on the Y chromosome of the gibbon and mouse (Nanda et al. 1990). Microsatellites also occur frequently in plant genomes and show extensive polymorphism (Akkaya et al. 1992; Morgante and Olivieri 1993; Lagercrantz et al. 1993). In *Arabidopsis thaliana* SSRs from the classes GA, CA, AT and A show random distribution and have been used to integrate the linkage map of this species (Bell and Ecker 1994). GATA- and GACA-probes have been reported to be suitable for cultivar identification in tomato (Vosman et al. 1992).

In the study reported here, a molecular linkage map was constructed from a *L. esculentum* × *L. pimpinellifolium* interspecific backcross population. This is the first complete linkage map reported for *L. pimpinellifolium*, one of the closest relatives of the cultivated tomato. The map covers the entire tomato genome at intervals averaging 10.7 cM and is comprised of RFLP, RAPD and morphological markers. The *L. pimpinellifolium* backcross map is compared with maps derived from other interspecific crosses with respect to marker segregation, marker order and chromosomal distribution of recombination. The map distribution of RFLP markers with RAPD and GATA microsatellite markers is also compared. The map and subsequent populations derived from this interspecific cross served as the basis for comparative QTL mapping across generations (Grandillo and Tanksley 1996; Tanksley et al. 1996).

Materials and methods

Plant material

The processing inbred *Lycopersicon esculentum* cv'M82-1-7' (denoted Ea) was crossed as pistillate parent to the accession LA1589 of the red-fruited wild species *L. pimpinellifolium* (denoted PM) originating in Peru. To obtain the BC_1 population, a single interspecific F_1 plant was backcrossed to the related processing inbred *L. esculentum* cv'E6203'(denoted Eb) using the F_1 as the female parent. The BC_1 plants were sown in flats in the greenhouse in April 1992. At the end of May, 264 BC_1 progeny were transplanted to the field in Ithaca, N.Y., and 257 were used to construct the genetic linkage map.

Morphological markers

The BC_1 population segregated for determinacy and uniform ripening – two simply inherited morphological traits controlled respectively by the *sp* and *u* genes (Tanksley et al. 1992). These two morphological markers were scored on the entire population and were included in the linkage map used for QTL analysis.

Molecular markers

RFLP analysis

Survey blots were prepared for the parents of the cross (Ea, Eb and PM) using DNA extracted from leaves as described by Bernatzky and Tanksley (1986a) except that mercaptoethanol was substituted for by sodium bisulfite. The DNA was digested with seven restriction enzymes (*Bst*NI, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Sca*I and *Xba*I) and subjected to Southern analysis following the procedures reported in Bernatzky and Tanksley (1986a). Four hundred and thirty DNA clones (cDNA and genomic) from the tomato high-density molecular map (Tanksley et al. 1992) were used for the survey. Probes were labeled with [^{32}P]-dCTP by primer extension (Feinberg and Vogelstein 1983). Hybridization and autoradiography were as reported by Bernatzky and Tanksley (1986b). A total of 115 RFLPs were chosen to construct the map used for QTL analysis. The BC_1 plants (257 in all) were analyzed using the same procedures described above.

RAPD analysis

Random amplified polymorphic DNAs (RAPDs) were tested as additional genetic markers (Williams et al. 1990). The 10-mers used as random primers in the PCR were purchased from Operon Technologies, Calif. The nucleotide sequence of the primers described in this paper are shown in Table 1. RAPD loci were named by OP followed by the Operon primer kit designation (a letter and a number); the size of each PCR product is not given. If more than one independently segregating polymorphic band were amplified by the same primer, capital letters were used to differentiate them – an A indicates the band of highest molecular weight.

The reactions were prepared as described by Martin et al. (1991) and were amplified using either a Perkin-Elmer Cetus DNA Thermal Cycler or a MJ Research PTC100 Programmable Thermal Controller. Reaction products were analyzed by electrophoresis on a 2% agarose gel and stained with ethidium bromide. A total of 500 primers were surveyed on the parental lines, with 26% of the primers showing reproducible and clearly scorable polymorphism (presence of the band in the PM parent and absence in both Ea and Eb parents). Of these primers 47 were assayed on the first 26 BC_1 plants for approximate localization on the genetic map. For 3 primers (OPAI-2, OPK-8 and OPV-20) the analysis was extended to the entire BC_1 population. These 3 primers were included in the map used for QTL mapping.

GATA microsatellite analysis

The oligonucleotide probe $GATA_8$ synthesized at Cornell was used in order to detect GATA-containing microsatellite loci valuable for mapping purposes (Vosman et al. 1992). Genomic DNA for the first 54 BC_1 plants was digested with *Taq*I, loaded on 0.9% agarose gels (20 × 24.5 cm) and electrophoresed for 48 h (25V) for the entire length of the gel. Southern blot analysis and hybridization were performed as described by Bernatzky and Tanksley (1986a, b). The oligonucleotide probe was [^{32}P]-labeled by primer extension (Feinberg and Vogelstein 1983). Bands and subsequent loci are designated by GATA followed by a letter. The band of highest molecular weight is indicated with an A.

Statistical analysis

Segregation ratios for marker classes were summarized and checked for conformity with the expected 1:1 ratios with the chi-square test using the computer programs MAP MANAGER (Manly 1993) and qGENE (Nelson 1994).

Table 1 List of 10-mer primers (Operon Technologies)

Primer	Sequence (5′ to 3′)	RAPD Locus	Chromosome number
OPAA-1	AGACGGCTCC	OPAA-1A	1
OPAC-6	CCAGAACGGA	OPAC-6	1
OPK-17	CCCAGCTGTG	OPK-17	1
OPQ-6	GAGCGCCTTG	OPQ-6	1
OPS-10	ACCGTTCCAG	OPS-10	1
OPT-17	CCAACGTCGT	OPT-17	1
OPV-3	CTCCCTGCAA	OPV-3	1
OPC-15	GACGGATCAG	OPC-15	1
OPAG-8	AAGAGCCCTC	OPAG-8B	2
OPAN-1	ACTCCACGTC	OPAN-1	2
OPB-8	GTCCACACGG	OPB-8	2
OPK-18	CCTAGTCGAG	OPK-18	2
OPG-12	CAGCTCACGA	OPG-12B	3
OPAI-2	AGCCGTTCAG	OPAI-2	5
OPAN-18	TGTCCTGCGT	OPAN-18	5
OPY-16	GGGCCAATGT	OPY-16	6
OPAA-7	CTACGCTCAC	OPAA-7A	6
OPAD-9	TCGCTTCTCC	OPAD-9B	6
OPAN-16	GTGTCGAGTC	OPAN-16	6
OPX-11	GGAGCCTCAG	OPX-11	7
OPY-4	GGCTGCAATG	OPY-4	7
OPAD-9	TCGCTTCTCC	OPAD-9A	7
OPAF-18	GTGTCCCTCT	OPAF-18	7
OPAM-6	CTCGGGATGT	OPAM-6	7
OPAM-9	TGCCGGTTCA	OPAM-9	7
OPK-8	GAACACTGGG	OPK-8	7
OPQ-5	CCGCGTCTTG	OPQ-5	7
OPT-16	GGTGTAACGCT	OPT-16	7
OPAA-1	AGACGGCTCC	OPAA-1B	8
OPAG-8	AAGAGCCCTC	OPAG-8A	8
OPAN-10	CTGTGTGCTC	OPAN-10	8
OPAN-14	AGCCGGGTAA	OPAN-14	8
OPF-20	GGTCTAGAGG	OPF-20	8
OPH-14	ACCAGGTTGG	OPH-14	8
OPS-20	TCTGGACGGA	OPS-20	9
OPG-12	CAGCTCACGA	OPG-12A	9
OPAD-2	CTGAACCGCT	OPAD-2	10
OPAJ-1	ACGGGTCAGA	OPAJ-1	10
OPV-14	AGATCCCGCC	OPV-14	10
OPV-20	CAGCATGGTC	OPV-20	10
OPA-2	TGCCGAGCTG	OPA-2	10
OPY-14	GGTCGATCTG	OPY-14	11
OPK-11	AATGCCCCAG	OPK-11	11
OPV-9	TGTACCCGTC	OPV-9	11
OPAC-2	GTCGTCGTCT	OPAC-2	12
OPA-17	GACCGCTTGT	OPA-17	12
OPF-16	GGAGTACTGG	OPF-16	12

Linkage analysis of the 257 BC_1 plants was performed with the software package MAPMAKER (Lander et al. 1987). In order to include a locus in a linkage group, a LOD threshold of 3.0 and a recombination fraction of at most 0.35 were chosen in the two-point analyses. The "order" and "ripple" commands were then used, respectively, to establish and verify the framework order of markers within groups. Markers were included within the framework map only if the LOD value for the ripple was greater than 3. Once the linear arrangement of marker loci along the chromosome was determined, the recombination frequencies between markers were estimated with multipoint analyses. The Kosambi mapping function (Kosambi 1944) was used to convert recombination frequencies to map distances in centiMorgans (cM).

The distribution of percentage of the recurrent parent (E) genotype in the 257 BC_1 progeny was estimated on the basis of marker genotypes and map distances using the computer program HYPERGENE (Young and Tanksley 1989).

The average donor segment length for the BC_1 population was calculated using the program qGENE (Nelson 1994) as described in Tanksley and Nelson (1996) and in Tanksley et al. (1996).

Results and discussion

Segregation and linkage analyses

A total of 120 genetic markers (115 RFLP, 3 RAPD and 2 morphological loci, the *sp* and *u* genes) were scored for

each of the 257 BC_1 plants. The resulting linkage map spanned 1279 cM with an average interval length of 10.7 cM between markers (Fig. 1). Due to the low level of polymorphism, a few regions of the genome (on chromosomes 7, 9 and 12) contained some intervals considerably longer than 20 cM.

The *L. esculentum* × *L. pimpinellifolium* (denoted E × PM) map obtained in this study was compared with the previously reported *L. esculentum* × *L. pennellii* (denoted E × P) high-density molecular map based on 1030 markers covering 1276 map units (Tanksley et al. 1992). The order of markers in the two maps is in good agreement. The E × PM map is based on a larger number of individuals and provides a better resolution of markers in a few instances, including (TG70-CT149) on chromosome 1, (TG525-TG129) on chromosome 3, (TG182-TG147) on chromosome 11 and (TG565-TG111) on chromosome 12 (Fig. 1).

Of the 120 markers scored in the E × PM population, 10 (8.3%) showed significant deviation from the expected 1:1 segregation ratio at $P < 0.05$ (Table 2; Fig. 1). Eight of these 10 markers displayed a higher frequency of heterozygotes than expected. These 8 markers mapped to chromosome 1 (2 markers), chromosome 8 (1 marker) and chromosome 11 (5 markers). The 2 loci skewed toward the E allele mapped to chromosome 5.

Segregation distortion has been reported in many interspecific crosses (Stephens 1949; Zamir et al. 1982; Wendel et al. 1987; Bonierbale et al. 1988; Paterson et al. 1988, 1991). Aberrant segregation might be caused by self-incompatibility genes, gametophytic selection and partly by viability selection of segregating plants (Zamir et al. 1982; Gebhardt et al. 1991).

When more distantly related wild species of tomato are used to generate segregating populations, more severe segregation distortion is often detected. In a BC_1 population with *L. chmielewskii* (CL), Paterson et al. (1988) reported skewness from the expected 1:1 segregation ratios for 69% of the markers; this constituted 21 distinct regions spread over all the 12 chromosomes.

Table 2 Significant deviation from the expected 1:1 ratio in the *L. esculentum* × *L. pimpinellifolium* BC_1 generation (*E esculentum* allele, *PM pimpinellfolium* allele)

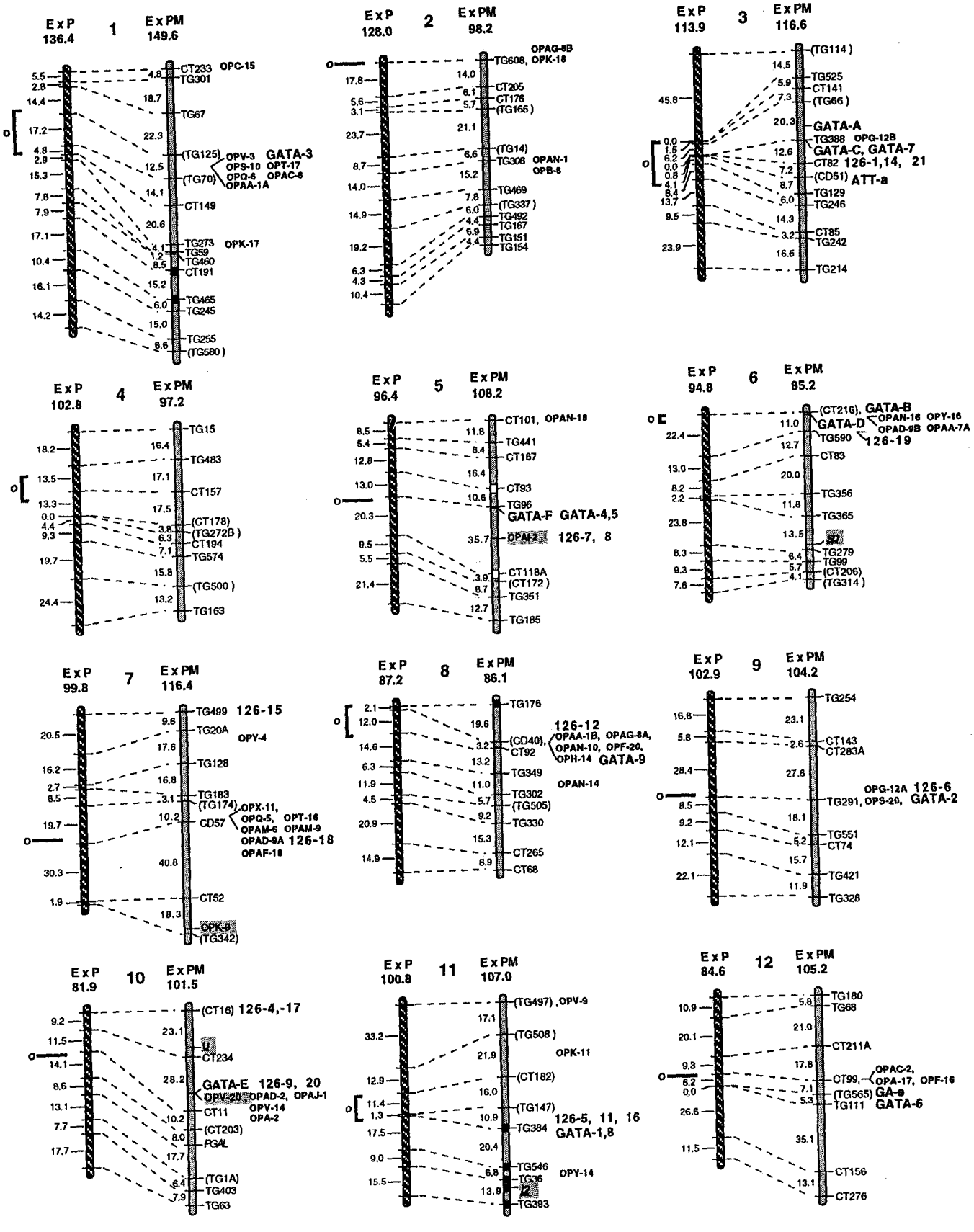
Locus	Chromosome	Genotypes	
		E/E	E/PM
CT191*	1	105	137
TG465*	1	110	142
CT93*	5	142	110
CT118A*	5	132	101
TG176**	8	96	156
TG384*	11	109	145
TG546**	11	102	148
TG36*	11	105	145
I2**	11	99	144
TG393**	11	106	150

*$P < 0.05$; **$P < 0.01$

Fig. 1 Genetic linkage maps of tomato derived from the *L. esculentum* × *L. pimpinellifolium* (E × PM) BC_1 population of 257 plants (*gray chromosomes*) and from the *L. esculentum* × *L. pennellii* (E × P) F_2 population of 67 plants (*hatched chromosomes*) (Tanksley et al. 1992). The order of the markers is the one obtained in the E × PM map. *Numbers on the left* correspond to map distances between markers in cM (Kosambi function). Chromosome map distances refer to those obtained using the set of 114 common RFLP markers indicated by *dashed lines*. Loci names are listed at *right* of chromosomes. Markers enclosed in *parentheses* have been located to corresponding intervals of the high-density tomato molecular map with LOD < 3; *underlined* are loci for which the position on the high-density molecular tomato map was approximated from previously published maps; all other markers reported have been ordered with LOD > 3 (Tanksley et al. 1992). *Shading* indicates the 6 extra markers (*sp*, *u*, *I2*, OPAI-2, OPK-8 and OPV-20) included in the map used for QTL analysis. Markers *separated by commas* cosegregate. **E × P map:** *Squared brackets* and *black lines* indicate approximate position of centromeres (see text for details). **E × PM map** only: OP primers indicated by *tick marks* have been mapped using the entire BC_1 population; OP primers at *right* of chromosomes have been mapped using a subset of 26 BC_1 plants. *Black squares* on chromosomes 1, 8 and 11 indicate markers with segregation significantly skewed in favor of the PM allele. *White squares* on chromosome 5 indicate markers with a segregation significantly skewed in favor of the E allele. [*CD* tomato leaf cDNA; *CT* tomato epidermal cDNA; *TG* tomato genomic clones; *OP* RAPD primers; *sp* self pruning; *u* uniform ripening; *PGAL* polygalacturonidase; *I2* resistance to *Fusarium oxysporum* race 2; *GATA-A-F* GATA-containing loci mapped on 54 E × PM BC_1 plants; *GATA-1-8*, *ATT-a*, *Ga-3* microsatellite loci mapped on 67 E × P F_2 plants (Broun and Tanksley 1996); *126*- GATA-containing loci mapped on a different E × P F_2 population (Arens et al. 1995)]

Twelve of these regions showed excess of the heterozygote (E/CL). Less distortion (51% of markers) was reported in a F_2 population derived from a cross of tomato with the wild species *L. cheesmanii* (CM) (Paterson et al. 1991). For the same interspecific cross, E × CM, a study was conducted on 97 recombinant inbred lines (Paran et al. 1995). In this case 73% of the markers showed significant deviation from the expected 1:1 ratio between the two homozygous classes, and 98% of the deviating markers were skewed in favor of the *esculentum* alleles. In a *L. esculentum* × *L. pennellii* F_2 population 80% of the markers, representing 16 linked segments spread over all 12 chromosomes, deviated significantly from the expected 1:2:1 frequency, with 82% favoring the P allele (de Vicente and Tanksley 1993).

The E × PM population reported here has less overall skewing than most of the other interspecific crosses reported for tomato. The marker showing the strongest distortion (toward the PM allele) was TG176, which is located at the top of chromosome 8 (Table 2; Fig. 1). On the same chromosome, but at a lower position, two previous studies based on isozyme markers (Tanksley et al. 1982; Vallejos and Tanksley 1983) showed strong skewing at the *Aps-2* locus, but in favor of the *esculentum* homozygote. It has been proposed that the two green-fruited species (*L. hirsutum* and *L. pennelli*) share a common genetic factor near *Aps*-2 on chromosome 8 that is deleterious when introduced into a *L. esculentum* genetic background. These results are in agreement with


ExP 136.4 1 ExPM 149.6
5.5 2.6 14.4 17.2 4.8 2.9 15.3 7.8 7.9 17.1 10.4 16.1 14.2
4.8 18.7 22.3 12.5 14.1 20.6 4.1 1.2 8.5 15.2 6.0 15.0 6.6
CT233 OPC-15
TG301
TG67
(TG125) OPV-3 GATA-3
(TG70) OPS-10 OPT-17 OPQ-6 OPAC-6 OPAA-1A
CT149
TG273 OPK-17
TG59
TG460
CT191
TG465
TG245
TG255
(TG580)
ExP 128.0 2 ExPM 98.2
17.8 5.6 3.1 23.7 8.7 14.0 14.9 19.2 6.3 4.3 10.4
14.0 6.1 5.7 21.1 6.6 15.2 7.8 6.0 4.4 6.9 5.4
OPAG-8B
TG608, OPK-18
CT205
CT176
(TG165)
(TG14)
TG308 OPAN-1 OPB-6
TG469
(TG337)
TG492
TG167
TG151
TG154
ExP 113.9 3 ExPM 116.6
45.8 0.0 1.5 6.2 0.0 0.8 4.1 8.4 13.7 9.5 23.9
14.5 5.9 7.3 20.3 12.6 7.2 8.7 6.0 14.3 3.2 16.6
(TG114)
TG525
CT141
(TG66)
GATA-A
TG388 OPG-12B
GATA-C, GATA-7
CT82 126-1,14, 21
(CD51) ATT-a
TG129
TG246
CT85
TG242
TG214
ExP 102.8 4 ExPM 97.2
18.2 13.5 13.3 0.0 4.4 9.3 19.7 24.4
16.4 17.1 17.5 3.8 6.3 7.1 15.8 13.2
TG15
TG483
CT157
(CT178)
(TG272B)
CT194
TG574
(TG500)
TG163
ExP 96.4 5 ExPM 108.2
8.5 5.4 12.8 13.0 20.3 9.5 5.5 21.4
11.8 8.4 16.4 10.6 35.7 3.9 8.7 12.7
CT101, OPAN-18
TG441
CT167
CT93
TG96
GATA-F GATA-4,5
OPAF-2 126-7, 8
CT118A
(CT172)
TG351
TG185
ExP 94.8 6 ExPM 85.2
22.4 13.0 8.2 2.2 23.8 8.3 9.3 7.6
11.0 12.7 20.0 11.8 13.5 6.4 5.7 4.1
(CT216), GATA-B
GATA-D OPAN-16 OPY-16 OPAD-9B OPAA-7A
TG590 126-19
CT83
TG356
TG365
S2
TG279
TG99
(CT206)
(TG314)
ExP 99.8 7 ExPM 116.4
20.5 16.2 2.7 8.5 19.7 30.3 1.9
9.6 17.6 16.8 3.1 10.2 40.8 18.3
TG499 126-15
TG20A OPY-4
TG128
TG183
(TG174) OPX-11, OPQ-5, OPT-16 OPAM-6 OPAM-9 OPAD-9A 126-18 OPAF-18
CD57
CT52
OPK-9
(TG342)
ExP 87.2 8 ExPM 86.1
2.1 12.0 14.6 6.3 11.9 4.5 20.9 14.9
19.6 3.2 13.2 11.0 5.7 9.2 15.3 8.9
TG176
(CD40), CT92 126-12 OPAA-1B, OPAG-8A, OPAN-10, OPF-20, OPH-14 GATA-9
TG349 OPAN-14
TG302
(TG505)
TG330
CT265
CT68
ExP 102.9 9 ExPM 104.2
16.8 5.8 28.4 8.5 9.2 12.1 22.1
23.1 2.6 27.6 18.1 5.2 15.7 11.9
TG254
CT143
CT283A
OPG-12A 126-6
TG291, OPS-20, GATA-2
TG551
CT74
TG421
TG328
ExP 81.9 10 ExPM 101.5
9.2 11.5 14.1 8.6 13.1 7.7 17.7
23.1 28.2 10.2 8.0 17.7 6.4 7.9
(CT16) 126-4,-17
CT234
GATA-E 126-9, 20
OPV-20 OPAD-2, OPAJ-1 OPV-14 OPA-2
CT11
(CT203)
PGAL
(TG1A)
TG403
TG63
ExP 100.8 11 ExPM 107.0
33.2 12.9 11.4 1.3 17.5 9.0 15.5
17.1 21.9 16.0 10.9 20.4 6.8 13.9
(TG497), OPV-9
(TG508)
OPK-11
(CT182)
(TG147)
TG384 126-5, 11, 16 GATA-1,8
TG546 OPY-14
TG36
TG393
ExP 84.6 12 ExPM 105.2
10.9 20.1 9.3 6.2 0.0 26.6 11.5
5.8 21.0 17.8 7.1 5.3 35.1 13.1
TG180
TG68
CT211A
OPAC-2, OPA-17, OPF-16
CT99, (TG565) GA-e
TG111 GATA-6
CT156
CT276

those reported for an E × CL backcross population (Paterson et al. 1988) and for morphological markers (Rick 1972). In a E × P F_2 population skewness was reported for the same region of chromosome 8 but toward the wild allele (de Vicente and Tanksley 1993). Recently, Eshed and Zamir (1995) also observed distorted segregation associated with the *Aps-2* locus, and they suggested the skewness in favor of the E allele to be due to the elimination of male gametes carrying the *L. pennellii* allele. The BC_1 population analyzed in the present study was obtained using the F_1 as pistillate parent. This excludes male gamete elimination as the possible cause of the distorted segregation observed for the E × PM BC_1 population.

Genome composition of BC_1 individuals

On average, the BC_1 individuals were inferred to be homozygous for the E alleles throughout 74% of the genome, which is very close to the expected 75% for a backcross (Fig. 2). Values for individual plants ranged from 61% to 87%. The individual with 87% appears to carry 8 chromosomal fragments from PM that range from 4.8 to 82.7 cM in length. The average length of the *pimpinellifolium* segment for the BC_1 population was 49.7 cM.

Distribution of recombination along chromosomes

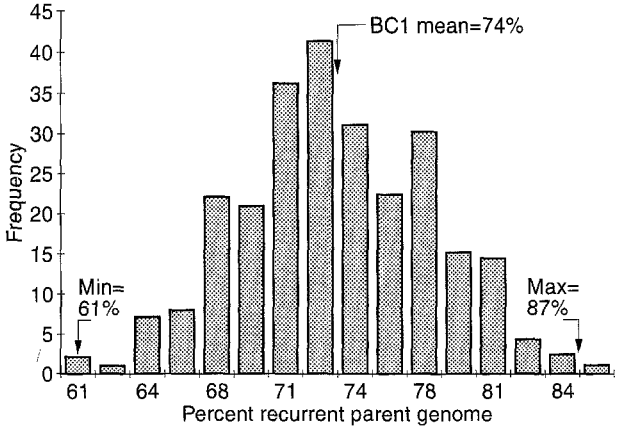
A set of 114 RFLP markers is shared between the E × PM population reported here and the E × P population used to create the high-density molecular linkage map of tomato (Tanksley et al. 1992) (Fig. 1). The total map distances, 1275 and 1230 cM respectively, were very similar for the two crosses. Considering the much closer phylogenetic relationship of *L. pimpinellifolium* to *L. esculentum* than that of *L. pennellii* to *L. esculentum*, an increase in total recombination might have been expected (Miller and Tanksley 1990). However, while there was no difference in the overall level of recombination in the two maps, there were considerable differences in the distribution of recombination. For example, the total length of chromosome 2 in the E × PM cross is 98 cM compared with 128 cM for the E × P cross. Similar but less pronounced trends were also seen with chromosomes 4 and 6. For most of the other chromosomes, the E × PM population revealed higher recombination frequencies.

On the E × P map, clusters of tightly linked markers were observed for several regions of the genome (Tanksley et al. 1992). Cytogenetic analysis indicated a high density of marker genes at the junction of the heterochromatin and euchromatin (Khush and Rick 1968). The dense clustering of markers that occurs mainly at centromeric and possibly telomeric regions has been attributed to suppressed recombination, which stems from a direct inhibitory effect on recombination of the centromere itself and/or adjacent centromeric heterochromatin (Tanksley et al. 1992). This hypothesis is supported by high-resolution genetic and physical mapping around the centromeres of chromosomes 7 and 9 (Ganal et al. 1989; Frary et al. 1996).

The approximate map position of the centromere is now known for most tomato chromosomes (Fig. 1). For the submetacentric chromosomes, 1 and 2, the centric positions have been obtained by RFLP mappng and by in situ hybridization of ribosomal sequences, 5S rDNA in chromosome 1 and 45S rDNA in chromosome 2 (Tanksley et al. 1988; Lapitan et al. 1991). The centromere of chromosome 3 has been recently located on the integrated molecular-classical map in the interval between the markers TG66 and TG246 (Koornneef et al. 1993; van der Biezen et al. 1994). By means of radiation-induced deletion mapping, Van Wordragen et al. (1994) placed the centromere of chromosome 6 in the GP79-*Aps1* interval of the integrated map. A more precise localization of the centromeres of chromosomes 7 and 9 has been achieved via dosage analysis of trisomic stocks (Frary et al. 1996). In addition, for chromosomes 5, 9, 10, 11 and 12, a good indication of the centric areas is provided by the positions of the respective breakpoints for the paracentric inversions that differentiate tomato and potato (Tanksley et al. 1992). Likely map locations for the centromeres of chromosomes 4 and 8 were predicted based on the relationship among the cytological, genetic and molecular tomato maps (Khush and Rick 1968; Tanksley et al. 1992). For most of the cases listed above, the proposed positions of the centromeres fall in regions of high marker density (Tanksley et al. 1992).

In the E × PM map, several of these putative centromeric regions (e.g. chromosomes 1, 3, 4, 5, 11 and 12) appear to be less suppressed in recombination than was observed in the E × P population (Tanksley et al. 1992) (Fig. 1). The most striking case is chromosome 3. The

Fig. 2 Distribution of percentage of recurrent parent (E) genotype in the BC_1 population, estimated on the basis of the marker genotypes and their relative distances



E × PM cross showed almost threefold more recombination in the centromeric interval TG66-TG246 than was observed in the E × P map. An approximately twofold increase in recombination frequency for the E × PM cross was observed for the centromeric intervals TG67-TG273 and CT211A-TG111 on chromosomes 1 and 12, respectively. The value went down to 1.5 times for chromosomes 4 (TG483-CT178), 5 (CT93-CT118A) and 11 (CT182-TG546).

The increased recombination frequency exhibited at several putative centric regions in the E × PM population versus the E × P population suggests that centric suppression of recombination may be greater in crosses with more distant species. This seems to agree with previous observations that, at least for some chromosomes, meiotic recombination is suppressed in the centromeric heterochromatin, and that this suppression can be more extreme in wide crosses (Khush and Rick 1967, 1968; Rick 1969, 1972). However, it is not an inviolable trend since a decrease in recombination was seen for centric regions corresponding to chromosomes 2, 6, 7 and 9. Two large F_2 populations derived from E × P and E × PM crosses have been analyzed with the goal of better resolution of the markers in the centromeric regions of chromosomes 7 and 9 (Frary et al. 1996). Significant suppression of recombination at the centric regions of these two chromosomes was found for both populations. High-resolution mapping extended to the centromeres of other tomato chromosomes should clarify the mechanism regulating the rate of chromosomal recombination in hetero- and euchromatic regions relative to different interspecific crosses.

GATA microsatellites and RAPDs are preferentially located near centromeres

The map positions of 47 RAPDs on the E × PM map appear to be non-randomly distributed. Thirty-five (75%) of them clustered around the putative centromeric regions (Fig. 1). For example, 6 primers mapped to the 13-cM TG125-TG70 centromeric interval on chromosome 1, 7 clustered close to the centric region on chromosome 7, 5 each were found in the centric areas of chromosomes 8 and 10 and 4 others mapped within a very short interval in the centromeric region of chromosome 6.

GATA microsatellites showed an even stronger preference for centromeres (Fig. 1). Six polymorphic bands from the *Taq*I digest could be scored clearly on the BC_1 progeny and were designated GATA-A to GATA-F. The bands GATA-A and GATA-C mapped on chromosome 3, precisely in the area where the E × P high-density molecular map showed a more than tenfold increase in marker density and where the centromere has been located (Koornneef et al. 1993; van der Biezen et al. 1994). On chromosome 5, the band GATA-F mapped in the region delimited by the markers CT93 and CT118A, which is also characterized by a high density of markers and where the position of the centromere has been inferred based on the tomato-potato comparative maps (Tanksley et al. 1992). An additional two bands, GATA-B and GATA-D, mapped at the top of chromosome 6, where the centromere has recently been mapped (Van Wordragen et al. 1994). The sixth band, GATA-E, mapped on chromosome 10, also in the region of high marker density and of inferred centric position (Tanksley et al. 1992).

The current study, in which a different interspecific population was used, supports the results of Arens et al. (1995), Broun and Tanksley (1996) and Y. Eshed and D. Zamir (personal communication) where different *L. esculentum* × *L. pennellii* F_2 populations were used to map GATA-containing loci. GATA clusters are thus now known to be located in most if not all of the centromeric regions of the tomato chromosomes.

The reason(s) for the clustering of GATA repeats near tomato centromeres is unknown. For certain types of SSRs possible functinal roles have been proposed; for example, in gene regulation or as hot spots for recombination (Tautz and Renz 1984; Walsh 1987). Therefore, one could speculate about a possible functional property of the SSRs related to the centromere itself. Alternatively, Charlesworth et al. (1986) suggest that regions of reduced recombination, such as those around centromeres, tend to accumulate and retain tandem repeats. Whatever the reason(s), it seems clear that both classes of markers, RAPD and GATA microsatellites, preferentially mark centromeres and are unlikely to provide uniform coverage of the tomato genome.

Conclusions

The genetic analysis conducted in this study shows that the E × PM map and the high-density molecular E × P map are very similar in total map distances and marker order. The mapping results obtained for the RAPDs and the GATA microsatellites suggest that these two classes of molecular markers are not randomly distributed but preferentially mark centromeric regions. Thus, maps constructed with only these types of markers are unlikely to provide the full genome coverage necessary for QTL mapping projects or marker-assisted selection in tomato.

A genetic linkage map has been constructed using 120 genetic markers (96% of which were represented by RFLPs) and covering the 12 tomato chromosomes at average intervals of 10.7 cM. This map provides the basis for QTL studies in *L. esculentum* × *L. pimpinellifolium* mapping populations (Grandillo and Tanksley 1996; Tanksley et al. 1996).

Acknowledgments We thank Drs. A. Frary and C. Nelson for their advice and helpful comments. This work was supported in part by grants from the National Research Initiative Coperative Grants Program, Plant Genome program USDA and by the Binational Agricultural Research and Development Fund (No. IS-1822-90C) to SDT. Partial funding was also received from the Cornell Center for

Advanced Technology (CAT) in Biotechnology, which is sponsored by the New York State Science and Technology Foundation and industrial partners. Silvana Grandillo was supported in part by the Italian CNR (National Research Council).

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131–1139

Arens P, Odinot P, van Heusden AW, Lindhout P, Vosman B (1995) GATA- and GACA-repeats are not evenly distributed throughout the tomato genome. Genome 38:84–90

Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258:1353–1355

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137–144

Bernatzky R, Tanksley SD (1986a) Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203:8–14

Bernatzky R, Tanksley SD (1986b) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103

Botstein D, White RL, Skolnick M, Davis R (1980) Construction of a genetic linkage map in man using restriction fragment polymorphisms. Am J Hum Genet 32:314–331

Broun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol Gen Genet 250:39–49

Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in citrus using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation responsive loci. Theor Appl Genet 89:606–614

Charlesworth B, Langley CH, Stephan W (1986) The evolution of restricted recombination and the accumulation of repeated DNA sequences. Genetics 112:947–962

Chittenden LM, Schertz KF, Lin YR, Wing RA, Paterson AH (1994) A detailed RFLP map of *Sorghum bicolor* × *S. propinquum*, suitable for high-density mapping, suggests ancestral duplication of Sorghum chromosomes or chromosomal segments. Theor Appl Genet 87:925–933

de Vicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134: 585–596

Dietrich W, Katz H, Lincoln SE, Shin H-S, Friedman J, Dracopoli NC, Lander ES (1992) A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423–447

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125

Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTLs. Genetics 141:1147–1162

Faure S, Noyer JL, Horry JP, Bakry F, Lanaud C, Gonzalez De Leon D (1993) A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). Theor Appl Genet 87:517–526

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Figdore SS, Kennard WC, Song KM, Slocum MK, Osborn TC (1988) Assessment of the degree of restriction fragment length polymorphism in *Brassica*. Theor Appl Genet 75:833–840

Frary A, Presting G, Tanksley SD (1996) Molecular mapping of the centromeres of tomato chromosomes 7 and 9. Mol Gen Genet (in press)

Ganal MW, Young ND, Tanksley SD (1989) Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol Gen Genet 215:395–400

Gawel NJ, Jarret RL, Whittemore AP (1992) Restriction fragment length polymorphism (RFLP)-based phylogenetic analysis of *Musa*. Theor Appl Genet 84:286–290

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homeologous tomato genome. Theor Appl Genet 83:49–57

Grandillo S, Tanksley SD (1996) QTL analysis of horticultural traits differenziating the cultivated tomato from the closely related species *Lycopersicon pimpinellifolium*. Theor Appl Genet 92:935–951

Kennard WC, Poetter K, Dijkhuizen A, Meglic V, Staub JE, Havey MJ (1994) Linkages among RFLP, RAPD, isozyme, disease-resistance, and morphological markers in narrow and wide crosses of cucumber. Theor Appl Genet 89:42–48

Khush GS, Rick CM (1967) Studies on the linkage map of chromosome 4 of the tomato and on the transmission of induced deficiencies. Genetica 38:74–94

Khush GS, Rick CM (1968) Cytogenetic analysis of the tomato genome by means of induced deficiencies. Chromosoma 23: 452–484

Koornneef M, Bade J, Hanhart C, Horsman K, Schel J, Soppe W, Verkerk R, Zabel P (1993) Characterization and mapping of a gene controlling shoot regeneration in tomato. Plant J 31: 131–141

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res. 21:1111–1115

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lapitan NLV, Ganal MW, Tanksley SD (1991) Organization of the 5S ribosomal RNA genes in the genome of tomato. Genome 34: 509–514

MacKill DJ, Salam MA, Wang ZY, Tanksley SD (1993) A major photoperiod-sensitive gene tagged with RFLP and isozyme markers in rice. Theor Appl Genet 85:536–540

Manly KF (1993) A Macintosh program for storage and analysis of experimental genetic mapping data. Mamm Genome 4:303–313

Martin GB, Williams JGK, Tanksley SD (1991) Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. Proc Natl Acad Sci USA 88:2336–2340

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistane in tomato. Science 262:1432–1436

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Nanda I, Deubelbeiss C, Guttenbach M, Epplen JT, Schmid M (1990) Heterogeneities in the distribution of $(GACA)_n$ simple repeats in the karyotypes of primates and mouse. Hum Genet 85:187–194

Nelson CJ (1994) Molecular mapping in bread wheat. PhD Thesis. Cornell University, Ithaca, N.Y.

O'Brien SJ (1993) Genetic maps. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Paran I, Goldman I, Tanksley SD, Zamir D (1995) Recombinant inbred lines for genetic mapping in tomato, Theor Appl Genet 90: 542–548

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinovitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127:181–197

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theor Appl Gent 90:503–510

Reinish AJ, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH (1994) A detailed RFLP map of cotton, *Gossipium hirsutum* × *Gossipium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847

Reiter RS, Williams JGK, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 89: 1477–1481

Rick CM (1969) Controlled introgression of chromosomes of *Solanum pennellii into Lycopersicon esculentum*: segregation and recombination. Genetics 62:753–768

Rick CM (1972) Further studies on segregation and recombination in backcross derivatives of a tomato species hybrid. Biol Zbl 91: 209–220

Schüeller C, Backes G, Fischbeck G, Jahoor A (1992) RFLP markers to identify the alleles on the *Mla* locus conferring powdery mildew resistance in barley. Theor Appl Genet 84:330–338

Stephens SG (1949) The cytogenetics of speciation in *Gossipium*. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34:627–637

Tanksley (1993) Mapping polygenes. Annu Rev Genet 27:205–233

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet (in press)

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49:11–25

Tanksley SD, Miller JC, Paterson AH, Bernatzky R (1988) Molecular mapping of plant chromosomes. In: Gustafson JP, Appels R (eds) Chromosome structure and function. Plenum Press, New York, pp 157–172

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND (1992) High-density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

Tanksley SD, Grandillo S, Fulton MT, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet (in press)

Tautz D (1989) Hypervariability of simple sequences as general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463–6471

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12: 4127–4138

Vallejos CE, Tanksley SD (1983) Segregation of isozyme markers and cold tolerance in an interspecific backcross of tomato. Theor Appl Genet 66:241–247

van der Biezen EA, Overduin B, Nijkamp HJJ, Hille J (1994) Integrated genetic map of tomato chromosome 3. Rep Tomato Genet Coop 44:8–10

Van Wordragen MF, Weide R, Liharska T, Van Der Steen A, Koornneef M, Zabel P (1994) Genetic and molecular organization of the short arm and pericentromeric region of tomato chromosome 6. Euphytica 79:169–174

Vosman B, Arens P, Rus-Kortekaas W, Smulders MJM (1992) Identification of highly polymorphic DNA regions in tomato. Theor Appl Genet 85:239–244

Walsh JB (1987) Persistence of tandem arrays: implications for satellite and simple-sequence DNAs. Genetics 115:553–567

Weber JJ, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801

Wendel JF, Edwards MD, Stuber CW (1987) Evidence for multilocus-genetic control of preferential fertilization in maize. Heredity 58: 297–301

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Yang X, Quiros CF (1995) Construction of a genetic linkage map in celery using DNA-based markers. Genome 38:36–44

Young ND, Tanksley SD (1989) Graphics-based whole genome selection using RFLPs. In: Helentjaris T, Burr B (eds) Development and application of molecular markers to problems in plant genetics. Current communications in molecular biology. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

Zamir D, Tanksley SD, Jones RA (1982) Haploid selection for low temperature tolerance of tomato pollen. Genetics 101: 129–137